The Comfstat – Automatically sensing thermal comfort for smart thermostats

Liliana Barrios
Department of Informatics
University of Zurich, Switzerland
Email: lilianaleonor.barriosaparicio@uzh.ch

Wilhelm Kleiminger
Dept. of Computer Science
ETH Zurich, Switzerland
Email: wilhelmk@ethz.ch

***Abstract*—Current smart thermostats aim to increase the efficiency of heating and cooling systems by adjusting the temperature whenever the conditioned zone becomes empty. However, targeting energy savings, these systems often fail to achieve a comfortable thermal environment for the inhabitants. We propose to increase thermal comfort by automatically monitoring the inhabitants' satisfaction with the thermal environment using commodity hardware. To this end we designed the Comfstat infrastructure and publish detailed temperature and heart-rate data of seven users of the system to the community. Using this data we show that thermal comfort can be inferred automatically from a combination of sensor data within 0.5 points on the ASHRAE scale.**

I. Introduction

A vital goal of current building automation systems is to reduce the energy consumption of heating, ventilation and air conditioning (HVAC) units. To this end, numerous projects in research [1], [2], [3] and industry [4], [5] have shown that it is feasible to build systems and algorithms to autonomously regulate the temperature in buildings in a "smart" way, for example based on actual occupancy.

However, non-technical challenges like sensor placement, uncertainty about the choices made by the system and doubts regarding achievable savings mean that user acceptance for such smart thermostats is still quite low [6], [7], [8]. And even if they show how energy consumption can be reduced, they often struggle with providing *thermal comfort* for the occupants. In fact, to save the maximum amount of energy, one could simply switch off the HVAC system altogether [9].

The strategy of most modern HVAC systems is to define a single comfort temperature. The building is then kept at this temperature throughout the day whenever it is occupied. While its value may have been obtained from experiments, the fixed temperature makes no allowances for the individual occupants' thermal preferences. In this paper, we target this problem by evaluating techniques to automatically sense thermal comfort from the occupants' heart rates as well as ubiquitous temperature and humidity sensors.

Thermal comfort has been described by the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) as "that condition of mind that expresses satisfaction with the thermal environment" [10]. As such it is defined by a combination of environmental as well as personal factors – including the clothing, health and mental state of the occupants. Since Ole Fanger first examined these factors in 1970 [11], further research has culminated various standards such as ASHRAE 55 [10] and ISO 7730 [12].

Following the observations made by Fanger, various researchers have sought to build systems to sense comfort automatically [13], [14], [15], [16]. As obtaining some of the parameters which influence thermal comfort (e.g., air speed and mean radiant temperature) is difficult in real-world settings, these systems combine sensory input (e.g., infrared temperature sensors) with user participation (i.e., through voting on the current comfort level) to enhance the accuracy of their models. While some of these systems include the metabolic rate in their calculations, none use available sensors to sense the heart rate as a proxy for the metabolic rate.

The importance of the metabolic rate for thermal comfort has recently been re-affirmed by Luo *et al.* [17]. The authors have shown that an increase in the metabolic rate is often a sign for discomfort. We build upon this work and the fact that the heart and the metabolic rate are closely related [18]. The Comfstat architecture allows for the collection of heart rate data from both Android Wear smartwatches and compatible Bluetooth chest straps. By using a machine learning approach and combining the heart rate with temperature and humidity data, we present a system that may deduce comfort automatically from the raw sensor data.

Besides showing that our approach can deduce thermal comfort with a mean error between 0.06 and 0.36 on the 7-point ASHRAE scale, we make the following contributions:

- An *infrastructure*[1] to collect sensory data from Android Wear watches, compatible BLE heart rate monitors, temperature and humidity sensors; as well as an application for registering ground truth thermal comfort through voting on both Android and Android Wear devices.
- A thermal comfort *data set*[2] comprised of seven participants, and including three different experiments.
- A comparison between heart rate data collected on current *smartwatches* and *dedicated chest straps*.
- An overview of the feature space and a detailed analysis of the performance of two approaches to predict thermal comfort.

[1] https://github.com/LilianaB/ComfstatInfrastructure
[2] https://github.com/LilianaB/ComfstatDataSet

The goal of this study is to provide a first analysis of using commodity hardware to automatically sense thermal comfort. To the best of our knowledge, the resulting data set, which we make publicly available, is the first of its kind. Our contributions are thus especially relevant for researchers interested in exploring the possibilities of using wearable technology to support thermal comfort prediction.

The outline of this paper is as follows. In Section II we introduce the main factors influencing thermal comfort and discuss related work. In Section III we show the components making up the Comfstat infrastructure. In Section IV we briefly describe the data set before we explain the feature extraction and selection in Section V. After we present our results in Section VI we conclude the paper in Section VII.

II. Background

ASHRAE 55 identifies six primary factors (metabolic rate, clothing level, air temperature, mean radiant temperature[3], air speed and humidity) that influence the occupants' satisfaction with the thermal environment. Based on climate chamber experiments, the standard provides equations to compute thermal comfort from these factors.

The metabolic rate thereby plays a crucial role. The metabolic heat production is determined by the energy balance of the human body [19]. Thermal balance is obtained when the heat loss to the environment is equal to the metabolic rate. The human body normally maintains a core temperature around 37 °C. When the *hypothalamus* (i.e., the portion of the brain controlling the core temperature) detects any changes in the surrounding conditions, it sends signals to the rest of the body to regulate the temperature by sweating or shivering. The person feels *uncomfortable*.

Following the 7-point scale first introduced by Fanger, the ASHRAE standard puts the thermal sensation of an individual on a scale ranging from $+3$ (hot) to -3 (cold). A value of 0 thereby indicates thermal neutrality and thus a comfortable environment. By combining the votes from multiple occupants, Fanger introduced the *Predicted Mean Vote* (PMV) and *Predicted Percentage Dissatisfied* (PPD) metrics. The PMV indicates the overall (average) satisfaction with the thermal environment. The PPD on the other hand shows the number of people who are feeling either too cold or too warm in the current environment.

a) Systems for sensing comfort: As both PMV and PPD do not explicitly track individual occupants' thermal comfort levels, Gao *et al.* introduced the "Predicted Personal Vote" model which extends the PMV for individual occupants [13], [14]. The model allows to achieve different micro-climates using personal heating and cooling appliances. Similarly to our approach, Gao *et al.* measure the temperature near the surface of the occupants' clothing using an infrared thermometer directed by a Microsoft Kinect camera. However, for the metabolic rate they depend on pre-computed tables.

[3]The average temperature of surfaces like the floor and the walls.

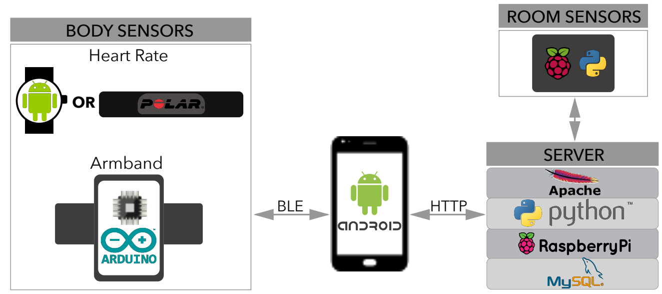

Fig. 1. Overview of the main components of the Comfstat sensing infrastructure.

Other work examines how occupants can dynamically vote on their current thermal comfort. In [15], Erickson *et al.* allow occupants to use their iPhone to vote on the 7-point ASHRAE scale, periodically. In [16], Lam *et al.* provide a mobile application for occupants to express their thermal satisfaction. We build on these ideas and extend it to smartwatches to gather ground truth for our Comfstat system.

Only few systems currently attempt to measure the metabolic rate to assess thermal comfort. Ciabattoni *et al.* propose a system very similar to ours, utilizing a Raspberry Pi to measure environmental parameters like the temperature and CO_2 levels [20]. In addition they use a smartwatch to measure the heart rate and skin temperature. However, while we learn individual comfort levels using user participation, the authors depend on fixed equations from ISO 8996 to deduce thermal comfort from the raw measurements [18].

Revel *et al.* propose an environmental monitoring system to monitor the PMV in an environment. However, while they discuss the importance of using the correct metabolic rate during the calculation they require the user to input their clothing level and metabolic rates manually through an Android device [21].

Abdallah *et al.* [22] employ an artificial neural network to link sensor data from wearable devices to thermal comfort. However, while their work seeks to approximate Fanger's equations, our model infers thermal comfort directly.

III. System Design

Comfstat's goal is to unobtrusively sense users' thermal comfort levels. We envision it to be installed as part of smart heating and cooling systems to better regulate setpoint temperatures. To achieve this, Comfstat is built around a mobile phone application supported by different sensors. Figure 1 shows the three key components of our architecture. *Sensors* worn on the body or placed in the room, a *smartphone* as sensing hub, and a *server* for data collection and analysis.

A. Server

As sensing and recording users' metabolic rates can be considered an intrusion into their privacy, one requirement of our system is to work standalone inside their home. However, in order to sense comfort, environmental parameters (e.g., air temperature) as well as personal factors (e.g., heart rate

Fig. 2. The mobile applications (left: smartphone, right: smartwatch) collect votes on the 7-point ASHRAE scale.

and thermal comfort votes) have to be combined. We thus aggregate the data on a local server. To this end we designed a Web application that allows for the storage and analysis to run on a low-cost single board system such as the Raspberry Pi. The Raspberry Pi runs a Apache HTTP Server and a MySQL database to store all user data for processing.

B. Mobile application on the smartphone

The mobile application serves as the central hub for gathering personal sensory data. It connects via Bluetooth Low Energy (BLE) to an Android Wear smartwatch and a chest-worn strap to collect heart rate data. It also obtains humidity and temperature values from sensors worn by the participants. All sensory values are quickly accessible to the user through the application.

The mobile application is realized on a Nexus 5 smartphone running the Android operating system. It synchronizes its data with the in-home server every five minutes. To preserve privacy and battery life, data is only collected and synchronized when the smartphone is situated within the home. This is achieved by validating that the user is connected to her home Wi-Fi.

In addition to the raw sensor values, the application allows us to collect ground truth thermal comfort data on the 7-point ASHRAE scale. A companion app on the smartwatch also allows for quick voting (cf., Figure 2). The data can be collected at frequencies between 5 and 15 minutes. Note that the collection of this "ground truth" data is only necessary for training the system. Once trained, the system should sense comfort automatically using the sensors introduced in the next section.

C. Sensors

The goal of Comfstat is to use sensors to automatically sense the users' comfort levels through a combination of body-worn and stationary sensors.

a) Heart rate: Heart rate (HR) measurements are a proxy for the metabolic rate. The metabolic rate is closely linked to thermal comfort as it determines how much thermal energy leaves the body [18]. Comfstat offers two modalities to reliably collect heart rate data.

First, a *smartwatch* application for the Android Wear operating system (running on an LG Watch R W110), which obtains heart rate values directly from an optical HR sensor built into the watch. The data are then transferred via the smartphone application to the server.

Secondly, heart rates can be collected directly through a dedicated sensor which implements the Bluetooth profile for heart rate service advertisement. We use an off-the-shelf Polar H7 heart rate sensor. As the Polar H7 uses a chest strap to measure the heart rate, its measurements are more accurate than those of the smartwatch.

In the following analysis, we will rely mainly on values from the Polar H7 strap to analyze the feasibility of automatically sensing comfort from the users' heart rates. Section VI-D gives an indication of the accuracy possible with current wrist-worn optical sensors.

b) Room temperature sensor: In order to collect the temperature in the conditioned area we use a DS18B20 temperature sensor connected to a Raspberry Pi. The temperature is retrieved every five minutes and forwarded to the in-home server.

c) Armband sensors: Normally, the air temperature varies slightly throughout the conditioned area (e.g., it will be slightly warmer near the windows due to the solar radiation). Therefore we provided users with an armband carrying an additional DHT22 temperature and humidity sensor. Note that the close proximity of the sensor to the body of the participant also means that its reading is likely to be influenced by the participant's body temperature. We will discuss this further in Section V when we elaborate on the features used to automatically deduce comfort.

The sensor is sampled every three seconds by an Adafruit Feather 32u4 Bluefruit LE board. The sensor data is exposed to the smartphone through the Bluetooth profile for environmental sensing. The board is powered by a Lithium Ion Polymer Battery with 2500 mAh, which allows for approximately five days of uninterrupted sensing.

IV. The Comfstat Data Set

The analysis presented in this paper is based on sensor data collected from seven participants in three controlled experiments. We refer to this data as the Comfstat data set and make it available to the research community.

Table I shows the recorded variables and their respective sampling interval. To achieve a higher accuracy for the HR measurements, participants were asked to use the Polar heart rate monitor instead of the smartwatch. Before each experiment, participants were asked to sign a consent form for volunteer subjects in an ergonomics investigation involving exposure to hot or cold temperatures. This consent form is based on the ISO 12894 standard [23].

The preliminary "cold" experiment was carried out with p_1 and p_2 only. Participants p_1, p_3, p_4, p_5, p_6 and p_7 took part in the main "controlled" temperature experiment. In a third experiment ("non-sedentary") with p_1 we assessed the effects of physical exercise on thermal comfort. Table III

TABLE I
COLLECTED VARIABLES.

Variable	Abbrev.	Interv.
Room temperature (°C)	`room_temp`	5 min
Temperature (armb.) (°C)	`ard_temp`	3 sec
Rel. humidity (armb.) (%)	`ard_hum`	3 sec
Heart rate (bpm)	`hr`	1 sec
Comfort (7-point scale)	`comfort`	1-5 min

TABLE II
PARTICIPANTS' PROFILES.

Participant	Gender	Age	Height	Weight
p_1	female	26	1.65 m	62 kg
p_2	male	28	1.80 m	66 kg
p_3	female	28	1.59 m	55 kg
p_4	female	23	1.61 m	67 kg
p_5	male	26	1.72 m	77 kg
p_6	female	28	1.64 m	65 kg
p_7	female	30	1.63 m	63 kg

summarizes which participants took part in which experiments. The following subsections explain each experimental setting.

A. Cold experiment

This preliminary experiment was conducted to test the Comfstat infrastructure and to measure the participants' responses to extreme thermal conditions. Two participants, p_1 and p_2 were subjected to low temperatures between 8 °C and 14 °C while wearing light clothing (i.e., 0.49 clo – 0.76 clo). The data was collected twice from 9.30 a.m. until 4.00 p.m. on two separate days. Figure 3 shows the timeline of each session.

During the preparation phase, participants were situated in a comfortable environment and explained how to use Comfstat. Next, the participants were asked to enter a cold (e.g., between 8 °C and 14 °C) environment for an hour. During this time, the participants were requested to provide feedback through the mobile application every one to five minutes. After one hour, the participants returned to their offices and resumed their normal daily routines (e.g., mainly sedentary activities). The room temperatures in p_1's office varied from 22 °C to 24 °C, while p_2's office temperature was stable at 22 °C. After their respective lunch breaks, the participants resumed their sedentary activity until 4 p.m.

B. Controlled temperature experiment

Six participants (p_1, p_3, p_4, p_5, p_6 and p_7) took part in the main experiment. Five of the participants are female and one is male. While the previous experiment served to measure the subjects' responses to extreme conditions and to test the system, the goal of this experiment was to measure the subjects' responses in more natural settings. For this purpose, participants were exposed to three different thermal environments: (i) *warm* (27 °C to 29 °C), (ii) *neutral* (23 °C), and (iii) *cold* (17 °C to 18 °C).

TABLE III
EXPERIMENT PARTICIPANTS.

Experiment	Participant
Cold	p_1, p_2
Controlled	p_1, p_3, p_4, p_5, p_6, p_7
Non-sedentary	p_1

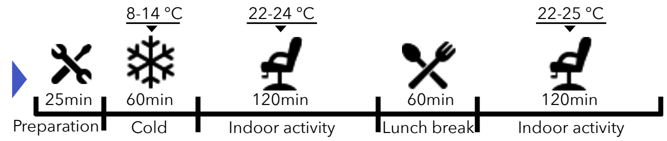

Fig. 3. Timeline of the "cold" temperature experiment.

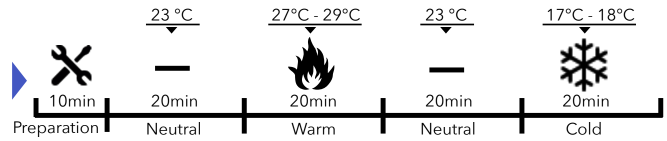

Fig. 4. Timeline of the "controlled" temperature experiment.

The main goal of this experiment is to analyze the subjects' body responses to temperature changes. To achieve the different thermal environments, three rooms were conditioned to the *warm*, *neutral*, and *cold* settings, respectively. Figure 4 shows the timeline of the experiment for each participant. The participants were exposed to each temperature setting for 30 minutes. In total, each experimental session lasted for 90 minutes. During the experimental period, all participants remained seated and only engaged in sedentary activities. Environmental data as well as the participants' satisfaction with the thermal environment were collected using the Comfstat infrastructure. Each participant undertook the experimental setting twice.

C. Non-sedentary experiment

The previous two experiments investigated how thermal comfort varies with different thermal environments when the subjects are sedentary. In order to analyze the effect of physical activity on thermal comfort, we conducted a separate 90-minute experiment with participant p_1. During the experiment p_1 wore light clothing (i.e., 0.44 clo) [24]. During the first 30 minutes, p_1 was asked to relax during a sedentary activity to obtain a baseline. Next, she was asked to perform physical exercise for 30 minutes followed by another 30-minute phase of relaxation.

D. Preliminary observations

Before going on to explain how we compute thermal comfort from the measured sensor values, we will identify a few key observations from the data set.

a) Thermal comfort is subjective: Figure 5 shows how the participants' thermal satisfaction is linked to the current temperature level. For each vote obtained through the smartphone application, Figure 5 shows the average temperature when the vote was cast. In general, participants show different thermal sensations at the same temperature levels, while females prefer higher temperatures. These results are confirmed by Karjalainen *et al.* [25], who found a significant impact

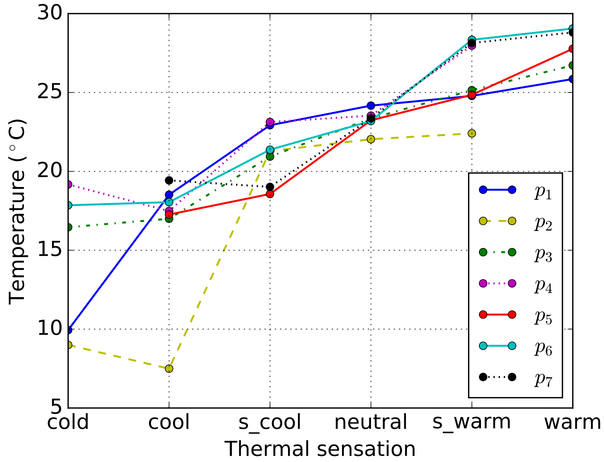

Fig. 5. Thermal comfort preference per user.

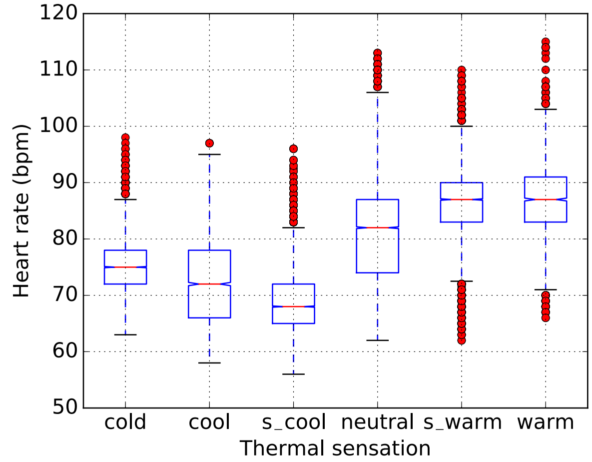

Fig. 6. P_1 heart rate per thermal comfort level.

TABLE IV
AVERAGE HEART RATE PER COMFORT LEVEL (-3: COLD, -2: COOL, -1: SLIGHTLY COOL, 0: NEUTRAL, 1: SLIGHTLY WARM, 2: WARM). NOTE THAT NO DATA FOR HOT (3) IS AVAILABLE.

	-3	**-2**	**-1**	**0**	**1**	**2**
(p_1)	75	72	69	81	86	87
(p_2)	84	84	77	93	94	-
(p_3)	74	67	72	74	70	71
(p_4)	84	80	84	85	84	-
(p_5)	-	79	76	79	81	80
(p_6)	79	80	81	78	84	83
(p_7)	-	69	68	72	65	68

TABLE V
PEARSON CORRELATION BETWEEN `ARD_TEMP` AND `COMFORT` ("CONTROLLED" TEMPERATURE EXPERIMENT).

	Pearson correlation					
	p_1	p_3	p_4	p_5	p_6	p_7
Session 1	0.85	0.80	0.81	0.79	0.76	0.78
Session 2	0.80	0.89	0.84	0.76	0.79	0.73

of gender on thermal comfort. Our results indicate that to automatically infer thermal comfort it may be necessary to create and train multiple models.

b) Heart rate higher in extreme environments: Figure 6 shows the distributions of the measured heart rates per thermal comfort level for p_1. The figure shows that the median heart rate (red line) tends to increase towards both cold and hot thermal sensations. Table IV summarizes these results for all participants. Our results are confirmed by Maohui Luo *et al.* [17], who found that the metabolic rate increases when the thermal sensation tends towards either end of the scale.

c) Thermal comfort indicated by temperature: During the preliminary analysis we also observed a correlation between participants' thermal comfort and the temperature measured by both the armband and the stationary sensor. Figure 7 shows, exemplary for p_1, how `ard_temp`, the temperature measured on the participants' arms, the room temperature `room_temp` and thermal satisfaction `comfort` change over the course of one "controlled" session. A summary of the correlation between `ard_temp` and `comfort` for all participants is shown in Table V.

V. MODEL

Our goal is to show how thermal comfort can be deduced automatically from the sensory input collected by the Comfstat infrastructure. According to the ASHRAE 55 standard, the satisfaction with the thermal environment can be expressed on a 7-point scale where -3 indicates *cold* and $+3$ indicates *hot*. This is also how thermal comfort is registered through the mobile application (cf. Figure 2).

Thus, thermal comfort may be considered a categorical variable. In Fanger's PMV and PPD calculations, however, fractional values are also possible. In the following we will therefore use two regression techniques – *linear regression* and *logistic regression*[4] to model the relationship between the measured data and thermal comfort. While the former outputs values on a continuous scale, the latter uses the discrete ASHRAE scale.

A. Feature extraction

Before we evaluate the performance of the two regression approaches in Section VI, we show how we identified and extracted a set of features which serve as good indicators for the comfort level. The features are based on the six fundamental factors (metabolic rate, clothing level, air temperature, mean radiant temperature, air speed and humidity) which define human thermal comfort [11].

1) Environmental factors: From the environmental factors (i.e., air temperature, mean radiant temperature, air speed and humidity), we only include air temperature and humidity by means of the `ard_temp`, `ard_humidity` and `room_temp` features. The first two are based on the sensor worn on the participant's arm while the latter is situated at a fixed position in the room.

[4] Regularization parameter set to 1 and not tuned for for best possible alpha.

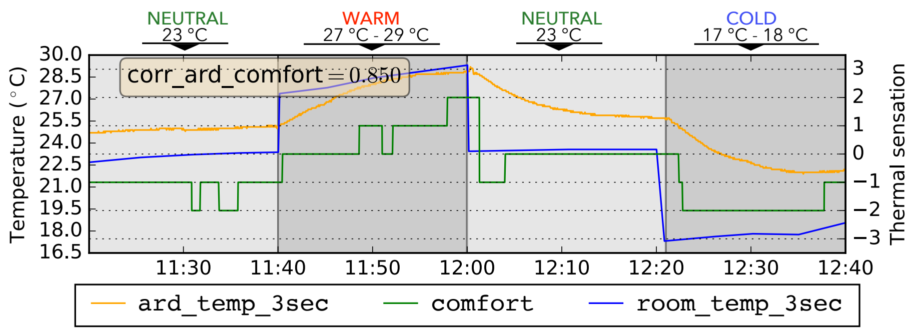

Fig. 7. P_1 Thermal comfort and temperature correlation.

TABLE VI
FEATURES COMPUTED ON COLLECTED DATA.

#	Feature name	Description
f_1	`polar`	Heart rate in bpm in 1s intervals
f_2	`room_temp`	Room temperature interpolated to 1s intervals
f_3	`ard_temp`	Armband temperature sensor interpolated to 1s internals
f_4	`ard_humidity`	Armband humidity sensor interpolated to 1s internals
f_5	`temp_delta`	Difference between room temperature and armband temperature interpolated to 1s intervals
f_6	`weight`	Participant's weight in kilograms
f_7	`height`	Participant's height in meters
f_8	`age`	Participant's age in years
f_9	`gender`	Participant's gender (1: female and 0: male)
f_{10}	`polar_1min_min`	Minimum heart rate (bpm) value registered in the last minute (sliding window)
f_{11}	`polar_1min_max`	Maximum heart rate (bpm) value registered in the last minute (sliding window)
f_{12}	`polar_1min_mean`	Mean heart rate (bpm) value registered in the last minute (moving average)
f_{13}	`room_temp_1min_min`	Minimum room temperature value registered in the last minute (sliding window)
f_{14}	`room_temp_1min_max`	Maximum room temperature value registered in the last minute (sliding window)
f_{15}	`room_temp_1min_mean`	Mean room temperature value registered in the last minute (moving average)
f_{16}	`ard_temp_1min_min`	Minimum armband temperature value registered in the last minute (sliding window)
f_{17}	`ard_temp_1min_max`	Maximum armband temperature value registered in the last minute (sliding window)
f_{18}	`ard_temp_1min_mean`	Mean armband temperature value registered in the last minute (moving average)
f_{19}	`ard_hum_1min_min`	Minimum armband humidity value registered in the last minute (sliding window)
f_{20}	`ard_hum_1min_max`	Maximum armband humidity value registered in the last minute (sliding window)
f_{21}	`ard_hum_1min_mean`	Mean armband humidity value registered in the last minute (moving average)

We do not model the air speed or the mean radiant temperature as both are difficult to measure outside the controlled environment of a climate chamber.

2) Metabolic rate: ISO 7730 establishes a linear relationship between the heart and metabolic rates of a person [12]. However, the calculation of the metabolic rate from the heart rate requires additional calibration to obtain parameters like the resting heart rate. To avoid this overhead, we decided to use the raw heart rate as a feature. In previous work, the heart rate has proved to be a good indicator for approximating energy expenditure [26]. We use the `polar` feature to denote the heart rate in bpm over a 1 s interval.

3) Clothing level and temperature differences: In addition to the heart rate, we model the difference `temp_delta` between the temperature measured on the body using the armband `ard_temp` and the temperature measured by the fixed sensor (i.e., `room_temp`). `temp_delta` is influenced by both, the room temperature and the participant's heat dissipation. As the armband is worn over the clothing, the weight of each temperature is influenced by the clothing level of the participant.

4) Temporal variations: Table VI summarizes the selected features. All features are computed at 1-second intervals. Whenever the granularity of the raw data was less than 1 s, linear interpolation was used. In addition to the raw data from the sensors (i.e., `polar`, `room_temp`, `ard_temp` and `ard_humidity`), we introduce, for each, their respective minimum, maximum and mean over the previous 60 seconds to capture temporal variations. We denote these by adding the suffixes `_1min_min`, `_1min_max` and `_1min_mean`.

Using the temporal variations on the `polar` feature (i.e., features f_{10}, f_{11} and f_{12}) our goal is to recognize when a person's activity level has increased for a short period of time (e.g., a person just climbed the stairs to reach her office). The thus increased metabolic rate may cause thermal discomfort only for a brief moment, making an adjustment to the thermal environment unnecessary.

Features f_{13} to f_{21} capture recent changes in the environmental conditions (i.e., temperature and humidity). This is important as the human body needs time to adapt to changes in the thermal environment. Figure 7 shows that when a participant entered a neutral from a warm environment, she needed several minutes before feeling comfortable.

5) Regression over multiple participants: Features f_6 to f_9 (e.g., `weight`, `height`, `age` and `gender`) are used specifically for evaluating the performance of the regression on different participants.

VI. Results

We investigate whether the satisfaction with the thermal environment – expressed on the 7-point ASHRAE scale by our participants – can be determined automatically from the raw sensor data. To this end, we tested both linear regression and logistic regression on the Comfstat data set. We will use the subscripts $_{\text{LIR}}$ and $_{\text{LOR}}$ to denote performance figures for the linear and logistic regression, respectively.

A. Evaluation

To evaluate the approaches, we merged all available data for each participant. The data was then shuffled and split into training and testing sets using 10-fold cross validation. For each fold, the regression was tested on $1/10$ of the data and trained on the remaining $9/10$. To ensure that each fold contained sufficient training data, the folds were chosen to preserve the distribution of samples for each thermal comfort category.

1) Metrics: We use four different metrics to measure the performance of the approaches. $\bar{e}$ and $\hat{e}$ denote the mean and median absolute error, respectively. These two metrics give an indication of how many points (on the ASHRAE scale) the comfort prediction is away from the actual sensation of the participants. The R^2 measure is included for reference in the figures as it is the standard means for determining the fit of a regression line.

As the median absolute error $\hat{e}$ is less useful for logistic regression – which outputs a categorical variable – we have also included the classification accuracy *Acc* as a metric. The accuracy gives the percentage of samples that have been classified correctly (e.g., the system correctly classified that a participant was feeling "cold" at a particular interval).

2) Baseline: As baseline we used three models. *Temperature only*, denoted by the subscript $_{\text{TEM}}$ is the performance obtained by using linear regression on the temperature data alone (*i.e.* only f_2 – `room_temp` – is used to predict comfort). The *Neutral* model (subscript $_{\text{NEU}}$) always assumes that a person is feeling comfortable (i.e., the vote on the ASHRAE scale is 0). This neutral vote is in the middle of the scale and has a maximum error of 2. The *Random* model (subscript $_{\text{RND}}$) assigns a uniform probability to all seven points on the ASHRAE scale and predicts a random comfort level at each interval. While the latter two approaches do not require any knowledge about the distribution of the comfort levels, they serve as lower bounds for the performance of the regression.

TABLE VII
Regression performance (controlled temperature experiment).

		p_1	p_3	p_4	p_5	p_6	p_7	**Avg.**
Linear R.	$\bar{e}_{\text{LIR}}$	0.46	0.53	0.47	0.39	0.43	0.25	0.42
	$\hat{e}_{\text{LIR}}$	0.38	0.47	0.42	0.32	0.37	0.16	0.35
Logistic R.	$\bar{e}_{\text{LOR}}$	0.36	0.28	0.10	0.11	0.27	0.06	0.20
	Acc_{LOR}	68%	75%	94%	89%	77%	94%	83%
Temp. only	$\bar{e}_{\text{TEM}}$	0.67	0.60	0.53	0.56	0.57	0.53	0.58
(Lin. Reg.)	$\hat{e}_{\text{TEM}}$	0.67	0.85	0.40	0.40	0.49	0.48	0.55
Neutral	$\bar{e}_{\text{NEU}}$	1.02	1.17	0.94	0.89	0.86	0.83	0.95
	Acc_{NEU}	33%	28%	49%	43%	36%	47%	39%
Random	$\bar{e}_{\text{RND}}$	1.95	2.05	2.02	1.92	1.93	1.91	1.96
	Acc_{RND}	15%	15%	14%	14%	14%	14%	14%

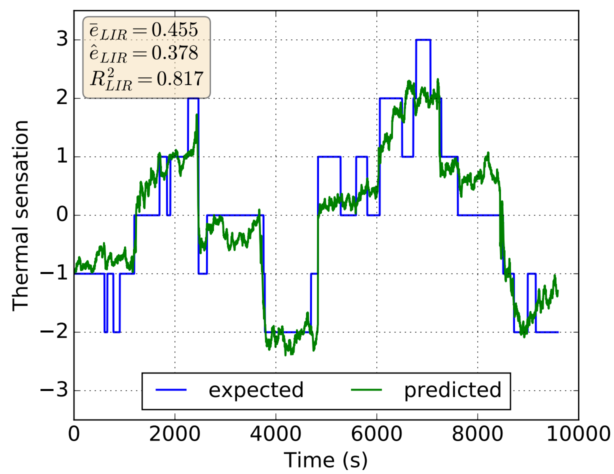

Fig. 8. P_1 Linear regression.

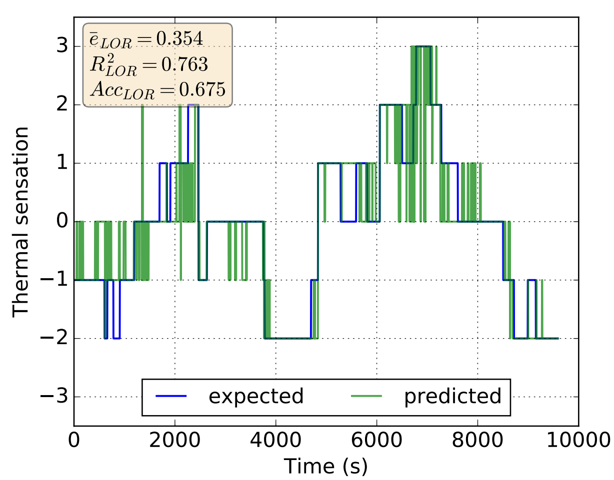

Fig. 9. P_1 Logistic regression.

B. Regression performance

Table VII shows the performance of the two regression approaches on the data from the *controlled temperature* experiment (cf. Section IV-B). Both regression approaches clearly outperform the baseline strategies. When only the temperature is used to deduce the participant's comfort levels, the mean error $\bar{e}_{\text{TEM}}$ is on average 0.58 points. The Neutral strategy has a mean error $\bar{e}_{\text{NEU}}$ between 0.83 and 1.17, while Acc_{NEU} is only 39% in the best case. As expected, the Random approach performs even worse with $\bar{e}_{\text{RND}}$ between 1.91 and 2.05 and an average Acc_{RND} of 14%.

For the linear regression, the mean error $\bar{e}_{\text{LIR}}$ varies between 0.25 and 0.53 for p_7 and p_3, respectively. This means, the prediction falls on average within 0.5 points of the actual thermal satisfaction of the participants. Similarly, the median error $\hat{e}_{\text{LIR}}$ ranges from 0.16 to 0.47, meaning that 50% of the time, the actual vote is within less than 0.5 points of the prediction. Using the full feature set means the mean error incurred (i.e., $\bar{e}_{\text{LIR}}$) is 0.16 points lower than the mean error incurred when only the temperature data is used (i.e., $\bar{e}_{\text{TEM}}$). The effect of the additional heart rate features is strongest for p_1 and p_7 where the respective differences in $\bar{e}$ are 0.21 and 0.28.

Figure 8 shows the results of the linear regression for p_1 over all data points. This data has been obtained by concatenating the results from all 10 testing folds. While it fails to notice short changes in the thermal sensation, the green regression line tracks the ground truth (expected) comfort level quite closely. As the output from the linear regression is continuous, it also seems[5] to capture intermediate comfort levels. However, as the thermal satisfaction is measured on the 7-point scale, these intermediate levels are deemed erroneous.

Thus, as the logistic regression models a categorical variable, one might expect it to show smaller errors. Indeed, Table VII shows that $\bar{e}_{\text{LOR}}$ varies between 0.06 for p_7 and 0.36 for p_1. The Accuracy Acc_{LOR} shows that between 68% (i.e., p_1) and 94% (i.e., p_4 and p_7) of intervals are classified correctly.

To understand why the logistic regression fails to achieve a higher accuracy for p_1, Figure 9 shows the result of the regression over the whole duration of the cold experiment. The uncertainty regarding the current comfort level results in a lot of fluctuations of the predicted thermal sensation. When the linear regression outputs a value between two distinct comfort levels (cf., Figure 8), the logistic regression often oscillates between them. While having a lesser effect on $\bar{e}_{\text{LOR}}$, this reduces the overall accuracy of the regression.

1) Discussion: The two regression approaches follow different goals. While the linear regression tries to model the relationship between the input variables (e.g., heart rate and temperature) and the output (i.e., thermal sensation) on a continuous scale, the logistic regression follows a classification (i.e., assigning input data to a number of distinct classes)

[5]The votes captured by the Comfstat infrastructure are expressed on the 7-point ASHRAE scale and thus do not allow for this granularity.

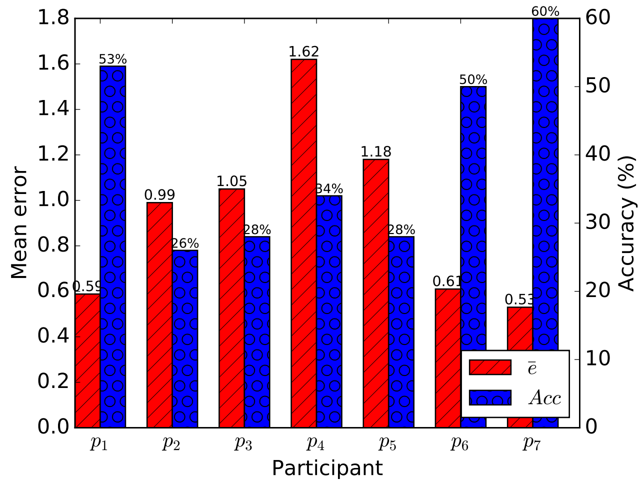

Fig. 10. Generalizability - Using other participants' data to predict comfort.

approach. Thus, both strategies have advantages and disadvantages. The linear regression results in a higher average error also because it may produce values outside the $[-3, 3]$ interval. On the other hand, it allows for detecting fractional comfort levels – something that the logistic regression is not capable of. The logistic regression may thus oscillate between two comfort levels as the available input data cannot be used to conclusively determine a single level.

These fluctuations are not good for heating and cooling systems as they may cause the system's setpoint temperature to fluctuate as well. In the worst case, this may lead to further oscillations breaking the control loop. In order to alleviate this problem, an additional smoothing step should be employed to reduce the number of fluctuations. A similar post-processing step could remove extreme values for the linear regression, effectively capping its output at -3 and 3.

However, the choice of method to predict thermal comfort also depends on how the individuals' thermal sensations are used. If the smart thermostat subsequently combines all votes from multiple inhabitants to compute a single assessment of the current thermal comfort level like Fanger's PMV, individual fluctuations may be less important. As our sample of seven participants is too small for such an analysis, we leave this question for future work.

C. Generalizability of trained model

During our preliminary analysis of the data set in Section IV-D we observed that thermal sensations varied significantly between participants for the same thermal environments. To understand how this affected the trained regression models, we used 5-fold cross validation. By leaving one participant out and training on all available data from the other participants we want to examine how well the regression generalizes over different participants. For this experiment, features f_6 to f_9 were introduced to model the weight, height, age and gender of the participants.

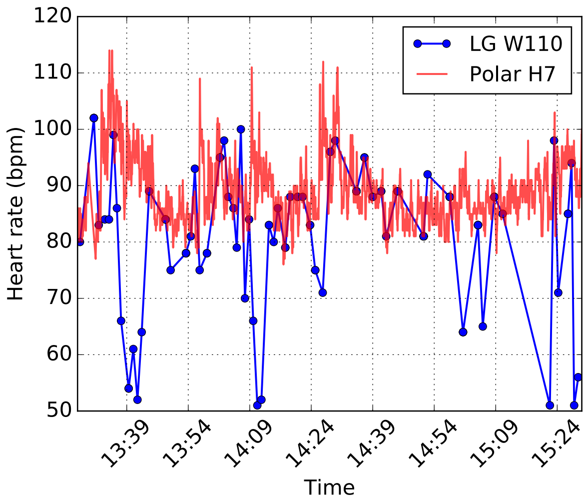

Fig. 11. HR monitoring accuracy (smartwatch vs chest strap). Pearson correlation: -0.045.

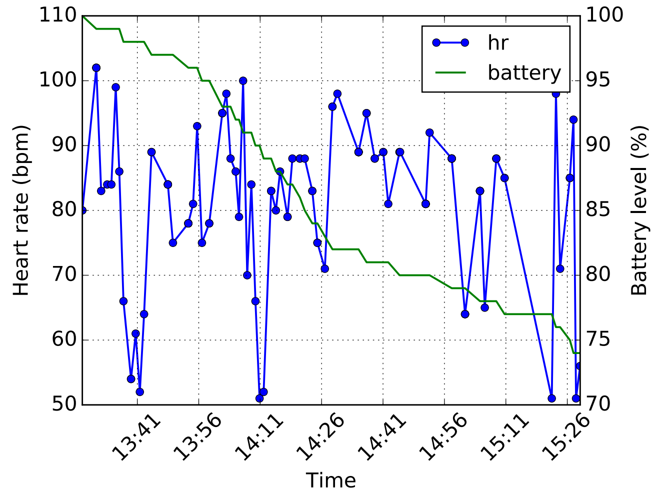

Fig. 12. Smartwatch battery drain.

Figure 10 shows the result of this experiment for the logistic regression. As before, $\bar{e}$ denotes the average absolute error, while *Acc* denotes the accuracy. For all participants $\bar{e}$ is above 0.5. Moreover, by using the other participants' data to predict p_2 to p_5, the logistic regression is wrong by a whole comfort level on average. This figure is similar to the performance of the Neutral baseline for both the controlled and cold temperature experiments. As the Neutral approach has no knowledge about any of the participants, thermal sensation does not seem to generalize for multiple people. This is important to note when considering alternative approaches like measuring the six primary factors defined in ASHRAE 55 and using equations derived from climate chamber experiments to deduce comfort [10]. By introducing a small training overhead that is made less strenuous through the use of smartwatches, our Comfstat approach offers personalized comfort prediction.

D. Replacing the chest strap

The major drawback of our current system is that it relies on a chest strap to measure the participants' heart rates. To understand whether current smartwatches might be a suitable substitute, we used the Comfstat infrastructure to collect heart rate data from a LG Watch R (W110) Android Wear smartwatch. In contrast to the Polar H7 chest strap which was sampled at 1 Hz, the smartwatch was only sampled every minute. Higher sampling frequencies were not possible as the watch would often time out as no value was detected.

Figures 11 and 12 show the result of a two-hour experiment during which we tested both the accuracy and the resulting battery drain. During this experiment, the smartwatch was only used to measure the heart rate. Notifications were disabled and the watch was not otherwise used.

Figure 11 shows that the heart rate measured by the smartwatch can deviate substantially from the chest strap, at times logging only half the value and well outside a reasonable range. This means that without significant thresholding and smoothing, the values from the smartwatch cannot currently be used to monitor comfort. The sampling interval is also restricted by the battery drain of the sensor. Over the course of the two-hour experiment, the smartwatch lost 25% of its capacity. This means that even if it was only used for sensing comfort, a smart thermostat could depend on the smartwatch for merely eight hours a day.

The low accuracy and restricted sampling interval thus make current smartwatches an unsuitable candidate for sensing the heart rate. However, as previous work has shown how pedometers and accelerometers can be used to monitor physical activity [27], future work might show how these can be integrated with the heart rate data to overcome periods of low accuracy and to reduce battery drain.

VII. Conclusions

We introduced our Comfstat infrastructure and showed how thermal comfort may be derived from participants' heart rates as well as environmental data including the room temperature and humidity. Furthermore, we made our data set publicly available to the research community. We show that high accuracies are possible when training a regression model using individual thermal sensation data and highlight that one cannot easily generalize the thermal sensation experienced by one participant in a particular environment to other participants. We propose to solve this problem by offering an easy to use calibration tool on both smartphones and smartwatches that allows occupants to periodically vote on the 7-point ASHRAE scale. Finally, we look into the future and examine the suitability of current smartwatches to measure the participants heart rates and conclude that further developments in sensor technology are necessary before they may be used to sense comfort.

REFERENCES

[1] M. C. Mozer, L. Vidmar, and R. H. Dodier, "The Neurothermostat: Predictive optimal control of residential heating systems," in *Advances in Neural Information Processing Systems*, M. C. Mozer, M. I. Jordan, and T. Petsche, Eds., vol. 9. The MIT Press, 1997, pp. 953–959.

[2] J. Lu, T. Sookoor, V. Srinivasan, G. Gao, B. Holben, J. Stankovic, E. Field, and K. Whitehouse, "The Smart Thermostat: Using occupancy sensors to save energy in homes," in *Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems (SenSys '10)*. Zurich, Switzerland: ACM, November 2010, pp. 211–224.

[3] J. Scott, A. J. B. Brush, J. Krumm, B. Meyers, M. Hazas, S. Hodges, and N. Villar, "Preheat: Controlling home heating using occupancy prediction," in *Proceedings of the 13th ACM International Conference on Ubiquitous Computing (UbiComp '11)*. Beijing, PRC: ACM, September 2011, pp. 281–290.

[4] "The Nest Learning Thermostat," http://www.nest.com. [Online]. Available: http://www.nest.com

[5] "tado° smart thermostat," http://www.tado.com/en. [Online]. Available: http://www.tado.com/en

[6] R. Yang and M. W. Newman, "Learning from a learning thermostat: Lessons for intelligent systems for the home," in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '13)*. Zurich, Switzerland: ACM, September 2013, pp. 93–102.

[7] W. Kleiminger, F. Mattern, and S. Santini, "Predicting household occupancy for smart heating control: A comparative performance analysis of state-of-the-art approaches," *Energy and Buildings*, vol. 85, pp. 493–505, December 2014.

[8] M. Pritoni, J. Woolley, and M. P. Modera, "Do occupancy-responsive learning thermostats save energy? A field study in university residence halls," *Energy and Buildings*, vol. 127, pp. 469–478, September 2016.

[9] G. Prins, "On condis and coolth," *Energy and Buildings*, vol. 18, pp. 251–258, December 1992.

[10] ASHRAE, "Thermal environmental conditions for human occupancy," American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, USA, ANSI/ASHRAE Standard 55-2010, 2010.

[11] P. O. Fanger, "Thermal comfort: Analysis and applications in environmental engineering," Ph.D. dissertation, Danmarks Tekniske Højskole, 1970.

[12] ISO, "Ergonomics of the thermal environment – analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices and local thermal comfort criteria," International Organization for Standardization, Geneva, Switzerland, ISO 7730:2005, 2005.

[13] P. X. Gao and S. Keshav, "SPOT: A smart personalized office thermal control system," in *Proceedings of the 4th International Conference on Future Energy Systems (e-Energy '13)*. Berkeley, CA, USA: ACM, May 2013, pp. 237–246.

[14] ——, "Optimal personal comfort management using SPOT+," in *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings (BuildSys '13)*. Rome, Italy: ACM, November 2013, pp. 1–8.

[15] V. L. Erickson and A. E. Cerpa, "Thermovote: Participatory sensing for efficient building HVAC conditioning," in *Proceedings of the 4th ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings (BuildSys '12)*. Toronto, ON, Canada: ACM, November 2012, pp. 9–16.

[16] A. H. yat Lan, Y. Yuan, and D. Wang, "An occupant-participatory approach for thermal comfort enhancement and energy conservation in buildings," in *Proceedings of the 5th International Conference on Future Energy Systems (e-Energy '14)*. Cambridge, United Kingdom: ACM, June 2014, pp. 133–143.

[17] M. Luo, X. Zhou, Y. Zhu, and J. Sundell, "Revisiting an overlooked parameter in thermal comfort studies, the metabolic rate," *Energy and Buildings*, vol. 118, pp. 152–159, April 2016.

[18] ISO, "Ergonomics of the thermal environment – determination of metabolic rate," International Organization for Standardization, Geneva, Switzerland, ISO 8996:2004, 2004.

[19] K. Parsons, *Human Thermal Environments: The Effects of Hot, Moderate, and Cold Environments on Human Health, Comfort, and Performance*, 3rd ed. FLorida, USA: CRC Press, April 2014.

[20] L. Ciabattoni, F. Ferracuti, G. Ippoliti, S. Longhi, and G. Turri, "IoT based indoor personal comfort levels monitoring," in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE'16)*, Las Vegas, NV, USA, January 2016.

[21] G. M. Revel, M. Arnesano, and F. Pietroni, "A low-cost sensor for real-time monitoring of indoor thermal comfort for ambient assisted living," in *Ambient Assisted Living: Italian Forum 2013*, S. Longhi, P. Siciliano, M. Germani, and A. Monteriù, Eds. Springer, 2014, pp. 3–12.

[22] M. Abdallah, C. Clevenger, T. Vu, and A. Nguyen, "Sensing occupant comfort using wearable technologies," *Construction Research Congress*, 2016.

[23] "Ergonomics of the thermal environment – medical supervision of individuals exposed to extreme hot or cold environments," International Organization for Standardization, Geneva, Switzerland, ISO 12894:2001, 2005.

[24] R. S. S. Z. Mansdorf and A. P. Neilsen, *Performance of Protective Clothing*, 2nd ed. Philadelphia, USA: ASTM, 1988.

[25] S. Karjalainen, "Gender differences in thermal comfort and use of thermostats in everyday thermal environments," *Building and Environment*, vol. 42, pp. 1594–1603, April 2007.

[26] S. M. Ceesay, A. M. Prentice, K. C. Day, P. R. Murgatroyd, G. R. Goldberg, W. Scott, and G. Spurr, "The use of heart rate monitoring in the estimation of energy expenditure: a validation study using indirect whole-body calorimetry," *British Journal of Nutrition*, vol. 61, pp. 175–186, April 1989.

[27] P. S. Freedson and K. Miller, "Objective monitoring of physical activity using motion sensors and heart rate," *Research Quarterly for Exercise and Sport*, vol. 71, pp. 21–29, July 2000.